

# Tail-Constraining Stochastic Linear-Quadratic Control: Large Deviation and Statistical Physics Approach

Michael Chertkov <sup>(1)</sup>, Igor Kolokolov <sup>(2,1)</sup>, and Vladimir Lebedev <sup>(2,1)</sup>

Center for Nonlinear Studies and Theoretical Division at LANL,  
& New Mexico Consortium, Los Alamos, NM, USA <sup>(1)</sup>

Landau Institute for Theoretical Physics, Moscow, Russia <sup>(2)</sup>

Standard definition of the stochastic Risk-Sensitive Linear-Quadratic (RS-LQ) control depends on the risk parameter, which is normally left to be set exogenously. We reconsider the classical approach and suggest two alternatives resolving the spurious freedom naturally. One approach consists in seeking for the minimum of the tail of the Probability Distribution Function (PDF) of the cost functional at some large fixed value. Another option suggests to minimize the expectation value of the cost functional under constraint on the value of the PDF tail. Under assumption of the resulting control stability, both problems are reduced to static optimizations over stationary control matrix. The solutions are illustrated on the examples of scalar and 1d chain (string) systems. Large Deviation self-similar asymptotic of the cost functional PDF is analyzed.

## I. INTRODUCTION

Stochastic differential equations are used both in control [1–7] and statistical physics [8–12] to state the problems. The two fields also use similar mathematical methods to analyze these equations. However, and in spite of the commonalities, there were relatively few overlaps between the disciplines in the past, even though the communications between two communities improved in the recent years. Some new areas in control, for example stochastic path integral control [13–16], have emerged influenced by analogies, intuition and advances in statistical/theoretical physics. Vice versa, many practical experimental problems in physics, chemistry and biology dealing with relatively small systems (polymers, membranes, etc), which are driven and experience significant thermal fluctuations, can now be analyzed and manipulated/controlled with accuracy and quality unheard of in the past, see for example [17, 18]. Besides, approaches from both control theory and statistical physics started to be applied to large natural and engineered networks, like chemical, bio-chemical and queuing networks [19–22]. Dynamics over these networks is described by stochastic differential equations, the networks have enough of control knobs, and they function under significant fluctuations which need to be controlled to prevent rare but potentially devastating failures. Related setting of stochastic optimization, i.e. optimization posed under uncertainty, has also came recently in the spot light of statistical physics inspired algorithms and approaches [23]. Convergence of these are related ideas motivated the manuscript, where we discuss analysis and control of rare events in the simplest possible, but practically rather widespread universal and general, linear setting. We realize that the general topic of linear control is well studied and many (if not all) possible questions, e.g. related to proper way of accounting for risk (rare events), were discussed in the field in the past. In spite of that, we still hope that this manuscript may also be useful not only to physicists, who may wish to explore new and largely unusual (in physics) formulations, but also to control theorists.

Consider first order (in time derivatives) stochastic linear dynamics of a vector  $x = (x_i | i = 1, \dots, N)$  over time interval  $t' \in [t; T]$

$$\frac{d}{dt'}x = Ax + Bu + \xi(t'), \quad (1.1)$$

where  $A$  and  $B$  are constant matrices;  $u(t')$  is the control vector applied at the moment of time  $t'$ ; and  $\{\xi\} = (\xi(t') | t' \in [t; T])$  is the zero mean, short-correlated noise with covariance  $V$

$$\langle \xi_i(t') \rangle = 0, \quad \langle \xi_i(t') \xi_j(t'') \rangle = \delta(t' - t'') V_{ij}, \quad i, j = 1, \dots, N \quad (1.2)$$

where one utilizes "statistical physics" notations for the expectation value (average) over noise,  $\langle \dots \rangle$ . Here in Eq. (1.2) and below the averages are over multiple possible realizations of the noise, each generating a new trajectory of the system,  $\{x\} = (x(t') | t' \in [t; T])$ , under given control  $\{u\} = (u(t') | t' \in [t; T])$ . The Eq. (1.1) is causal, thus assuming retarded (Stratonovich) regularization of the noise on the right-hand-side of the discrete version of Eq. (1.1). The physical meaning of the vectors and matrices in Eq. (1.1) is as follows.  $A$  is the matrix explaining stretching, shearing and rotation of the system trajectory in the  $N$ -dimensional space if the control and external noise would not be applied. Matrix  $B$  describes possible limitations on the degrees of freedom in the system one can control. To simplify notations we consider signal, control and noise vectors having the same dimension,  $N$ , where thus  $B$  is quadratic. The setting of Eqs. (1.1,1.2) is classic one in the control theory. It describes the so-called Linear-Quadratic (LQ)

stochastic control problem, which was introduced in [2, 3, 5, 6] and became foundational for the control theory as a field, see e.g. [24, 25] and references therein. In the classical formulation one seeks to solve the following optimization,  $t \in [0; T]$ :

$$\min_{\{u\}} \left\langle J(t; T; \{u\}, \{x\}) \right\rangle, \quad (1.3)$$

$$J(t; T; \{u\}, \{x\}) \equiv \frac{1}{2} x^*(T) F x(T) + \frac{1}{2} \int_t^T dt' (x^*(t') Q x(t') + u^*(t') R u(t')),$$

where  $Q, R$  and  $F$  are pre-defined stationary (time independent) symmetric positive matrices and one uses the superscript asterisk,  $*$ , to mark transposition.  $J(t; T; \{u\}, \{x\})$  (later on, and when it is not confusing, we will use the shortcut notation  $J$ ) is a scalar quadratic cost functional of the state vector  $\{x\} = (x(t')|t' \in [t; T])$ , and the control vector,  $\{u\} = (u(t')|t' \in [t; T])$  evaluated for all intermediate times  $t'$  from the  $[0; T]$  interval. Here in Eqs. (1.3) (and everywhere below in the manuscript) the average over noise  $\{\xi\}$  includes conditioning to Eq. (1.3), i.e.  $\{x\}$  is dependent on realization of the noise,  $\{\xi\}$ , and on the control,  $\{u\}$ , according to Eq. (1.1). It is assumed that the stochastic LQ control is evaluated off-line, i.e. the optimal solution  $u_*(t; x(t))$  of Eq. (1.3) is computed and saved prior to executing actual experiment for any initial condition  $x(t)$  at any  $t$ . Then in the course of the actual experiment (execution of the dynamics)  $x(t)$  is measured at any time  $t$  and respective  $u_*(t; x(t))$  is applied. (When observation of  $x(t)$  is partial and noisy one needs to generalize the stochastic LQ control, for example considering the stochastic Linear Quadratic Gaussian (LQG) control, see e.g. [25] for details.) We also assume (and the details will be clarified below) that the optimal control succeeds, i.e. the systems stabilizes and  $J$  does not grow with  $T$  faster than linearly.

An unfortunate caveat of the LQ setting (1.3) is in the lack of fluctuations control: even though the LQ solution is optimal in terms of minimizing the expectation value of the cost functional it may generate very significant fluctuations when it comes to analysis of the  $\mathcal{J} \gg \langle J \rangle$  tail of the Probability Distribution Function,  $\mathcal{P}(\mathcal{J})$ , of the cumulative cost  $\mathcal{J}(T; x(0)) \equiv J(0; T; \{u_*\}, \{x\})$ . Stochastic Risk Sensitive LQ (RS-LQ) scheme [26–28] was introduced to improve control of the abnormal fluctuations of  $J$ . RS-LQ constitutes the following generalization of the LQ scheme (1.3)

$$\max_{\{u\}} \langle \exp(-\theta J) \rangle, \quad (1.4)$$

where  $\theta$  is a pre-defined parameter. Intuitively one relates the case of positive  $\theta$  to a risk-avert optimum. It is assumed within the standard RS-LQ scheme that  $\theta$  is fine tuned by some additional considerations. Note that, as shown in [29], the RS-LQ control is also ultimately related to the so-called  $H_\infty$ -norm robust control. (See also [30] for further discussion of the relation.)

In this paper we analyze two natural modifications of the stochastic RS-LQ control. The two schemes can both be interpreted in terms of the RS-LQ approach supplemented by an additional optimization over  $\theta$ . Our first, *Tail-Optimum* (TO), scheme consists in the following modification of the LQ (1.3) and RS-LQ (1.4) ones

$$\min_{\{u\}} \mathcal{P}(J = j \cdot (T - t) | \{u\}). \quad (1.5)$$

In words, the TO-LQ control minimizes (at any time  $t$  and given the current observation  $x(t)$ ) the probability of the current value of the cost functional  $J(t; T; \{u\}, \{x\})$  evaluated at a predefined value,  $j \cdot (T - t)$ , where thus  $j$  is the only external parameter left in the formulation. Another strategy, which we call *Chance-Constrained* LQ (CC-LQ), in reference to similar formulations in optimization theory [31–33], consists in minimizing the mean of the cost functional under condition that the tail probability evaluated at  $j \cdot (T - t)$  does not exceed the prescribed threshold value  $\varepsilon(t; T)$

$$\min_{\{u\}} \langle J \rangle \quad \text{s.t. } \mathcal{P}(J = j \cdot (T - t) | \{u\}) \leq \varepsilon(t; T). \quad (1.6)$$

Main objectives, and consequently results of this study, are

- To extend the asymptotic,  $T \rightarrow \infty$ , approach, developed in the past for LQ (1.3) and RS-LQ (1.4) optimal controls to the new TO-LQ (1.5) and CC-LQ (1.6) optimal settings. At  $T \rightarrow \infty$  the optimal control takes the following universal linear in  $x$  form

$$u_*(t; x) = -Kx, \quad (1.7)$$

where  $K$  is  $t$ -independent but model dependent matrix. The condition of the system stability, intuitively translating into the expectation that  $\mathcal{J}$  grows not faster than linearly with  $T$ , naturally requires that all the eigenvalues of the

stability matrix,  $\mu = BK - A$ , have positive real part. The linearity of the optimal control (1.7) in  $x$  is a direct consequence of the linearity of the initial dynamical system. Time-independence and initial condition-independence of the optimal control (1.7) are asymptotic: they are achieved at  $T \gg \tau_*$ , where  $\tau_*$  can be estimated as the inverse of the absolute value of the  $(BK - A)$ 's eigenvalue with the smallest real part. System of algebraic equations defining  $K$  implicitly for the TO-LQ and CC-LQ cases are presented and then juxtaposed against the previously analyzed cases of the LS and RS-LQ controls. (See Eqs. (4.1,4.3,4.5).) Finding optimal control is reduced to optimization over time-independent,  $K$ . The resulting dependencies are homogeneous in time, with  $t$  and  $T$  always enter in the  $T - t$  combination. (This also simplifies the analysis allowing to set  $t = 0$ .)

• *To analyze statistics of the optimal cost functional,  $\mathcal{J}$ , in the Large Deviation (LD) regime, i.e. at large but finite  $T$ .* We show that in the stable regime the PDF of  $\mathcal{J}$  attains the following universal LD form

$$\log \mathcal{P}(\mathcal{J}) \sim -TS(\mathcal{J}/T), \quad (1.8)$$

where the LD function,  $\mathcal{S}(j)$ , is a convex function of its argument found implicitly (in a closed algebraic form, which may or may not yield an efficient algorithm) for the four cases (of LS, RS-LQ, TO-LQ and CC-LQ controls) considered. The LD function shows a universal,  $\mathcal{S}(j) \rightarrow aj$ , tail at large (i.e. larger than typical)  $j$ , where the value of positive  $a$  depends on the model. This suggests, in particular, that it is natural to choose in the CC optimization (1.6),  $\varepsilon(t; T) = \exp(-c(T - t))$ , for the threshold, with  $c$  been a constant. To derive compact algebraic expressions for the LD function we, first, analyze the generating function of  $\mathcal{J}$  evaluated at linear  $u$  parameterized by  $K$  as in Eq. (1.7),

$$\mathcal{Z}(\theta; K) \equiv \langle \exp(-\theta \mathcal{J}) \rangle_*, \quad (1.9)$$

then express the optimization/control objective as a convolution of the integral or differential operator/kernel in  $\theta$  (the choice will depend on the model) and  $\mathcal{Z}(\theta; K)$ , and finally evaluate optimization over  $K$  in the asymptotic LD approximation. Here in Eq. (1.9) the low asterisk mark  $*$  in the expectation/average (over noise and constrained to Eq. (1.1)) indicates that the control vector is taken in the form of the linear ansatz,  $u \rightarrow -Kx$ , where  $K$  is left yet undetermined.

The remainder of the manuscript is organized as follows. We start discussing the deterministic case (of zero noise) in Section II. This regime is of interest for two reasons. First, in the asymptotic of zero noise the four, generally different, control schemes become equivalent. Besides, and as well known from the classical papers [2, 24, 25], optimal control in the bare LQ case (correspondent to minimization of the cost function average) is not sensitive (and thus independent of) the level of the noise. Section III is devoted to analysis of the generating function (1.9), the average value of the cost function and the tail of the cost function distribution restricted to yet unspecified value of  $K$ . Optimization over  $K$ , resulting in the known RS-LQ optimal relations and also derivation of the new optimal relations for  $K$  in the TO-LQ and CC-LQ cases, is discussed in Section IV. We describe and compare asymptotic Large Deviation forms of the cost function PDF,  $\mathcal{P}(\mathcal{J})$ , in the optimal regimes. In this and preceding Section we also discuss many times the illustrative "scalar" example, where  $x$  and  $u$  are scalars. An infinite system example, of a "string" formed from a linear 1d chain, is discussed in Section V. We conclude and discuss related future challenges in Section VI.

## II. DETERMINISTIC CASE AND LQ-OPTIMAL CONTROL

We start this Section from a disclaimer: all results reported here are classical, described in [2, 24–29] and latter papers and books, see e.g. [34–36]. We present it here only for making the whole story of the manuscript self-explanatory and coherent.

When the noise is ignored, Eq. (1.1) should be considered as a deterministic constraint, reducing any of the optimal control schemes (1.3,1.4,1.5,1.6) to a simple variation of the cost functional (1.4) over  $u$ . Using the standard variational technique with a time dependent Lagrangian multiplier for the constraint, and then excluding the multiplier one derives the equation

$$\frac{d}{dt'} u^* + u^* R B^{-1} A B R^{-1} = x^* Q B R^{-1}, \quad (2.1)$$

which should be supplied by the boundary condition (also following from the variation),  $u^*(T) + x^*(T) F B R^{-1} = 0$ . (Let us remind that we choose the notations where the dimensionality of  $u$  coincides with the dimensionality of  $x$ . We also assume that inverses of all the matrices involved in the formulation are well defined. This assumption is not critical and is made here only to simplify the notations. In the general case when some of the matrices, in particular  $R$ , are not full rank, one can generalize the formulas properly, using a proper notion of the pseudo-inverse.) Substituting,  $u = -R^{-1} B^* \Pi x$ , in Eq. (2.1) one arrives at the following equation for  $\Pi$

$$\frac{d}{dt'} \Pi + \Pi A + A^* \Pi - \Pi B R^{-1} B^* \Pi + Q = 0. \quad (2.2)$$

with the boundary condition  $\Pi(T) = F$ . Eq. (2.2), solved backwards in time, results in  $\Pi(t)$  and then,  $u_*(t; x) = -R^{-1}B^*\Pi(t)x = -Kx$ .

To gain a qualitative understanding of the backwards in time dynamics of  $\Pi$ , let us briefly discuss the simplest possible case with all the matrixes entering Eq. (2.2) replaced by scalars, then yielding the following analytic solution for the optimal  $K$

$$K = \frac{A - \sqrt{A^2 + \frac{QB^2}{R}} \left( \tanh \left( \sqrt{A^2 + \frac{QB^2}{R}}(t - T_0) \right) \right)^{\pm 1}}{B}, \quad (2.3)$$

where  $T_0$  and  $\pm 1$  are chosen to satisfy the boundary condition,  $K(T) = BF/R$ . When  $T \gg \tau = 1/\sqrt{A^2 + QB^2/R}$ , the backwards in time dynamics saturates (after a short  $\sim \tau$  transient) to a  $F$ -independent constant, resulting from replacing  $\tanh$  in Eq. (2.3) by  $-1$ . Therefore, in the stationary regime,  $T \rightarrow \infty$  the optimal control is with the constant in time, frozen  $K$ . One also finds that the optimal control in the one dimensional deterministic case is always stable,  $\mu = KB - A > 0$ .

Returning to the general (finite vector) case one concludes that when  $T$  is sufficiently large the optimal control is of the form described by Eq. (1.7), i.e. it is linear in  $x$  and asymptotically time independent, with  $K = R^{-1}B^*\Pi_0$  where  $\Pi_0$  solves Eq. (2.2) with the first term replaced by zero. It is well known in the control theory that (under some standard common sense assumptions on  $B$  and  $R$  matrixes) stable solution of the system of the algebraic Riccati equations is unique and moreover it can be found efficiently. (See e.g. Chapter 12 of Sec[37] and references therein.)

Let us now discuss the bare LQ control, now in the presence of the noise. Since Eq. (1.1) is linear, one can naturally split the full solution into a sum,  $x = x_1 + x_2$ , where  $x_1$  satisfies Eq. (1.1) without noise and it is equivalent to the noise-less solution, just discussed in this Section. Then, the second term satisfies,  $dx_2/dt' = Ax_2 + \xi$ , with  $x_2(t) = 0$ . However, since the noise is zero mean,  $\langle \xi \rangle = 0$ ,  $x_2$  is zero mean too, i.e.  $\langle x_2 \rangle = 0$ . Next, let us analyze the split of term in  $\langle J \rangle$ , which is the optimization objective of the LQ scheme. Since,  $x_1$  and  $x_2$  are independent (by construction) and because  $x_2$  is zero mean,  $\langle J \rangle$ , splits into two terms,  $\langle J_1 \rangle + \langle J_2 \rangle$ , each dependent on  $x_1$  and  $x_2$  vectors only.  $\langle J_1 \rangle$  is simply equivalent to  $J$  analyzed above in the deterministic case, while  $\langle J_2 \rangle$  is  $u$ -independent, thus not contributing the optimization at all. To summarize, the LQ optimal control is not sensitive to the noise and it is thus equivalent to the deterministic (noiseless) case described above.

### III. GENERATING FUNCTION

Consider the Generating Function (GF),  $\mathcal{Z}(\theta; K)$ , defined by Eq. (1.9).  $\mathcal{Z}(\theta; K)$  is of an obvious relevance to the RS-LQ scheme, but it is also useful for analysis of other schemes as well, because of the following (Laplace transform) relation to the PDF of  $\mathcal{J}$ :

$$\mathcal{Z}(\theta; K) = \int_0^\infty d\mathcal{J} \exp(-\theta\mathcal{J}) \mathcal{P}_*(\mathcal{J}), \quad (3.1)$$

where (as before) the asterisk in the sub-script indicates that the PDF was evaluated at  $u = Kx$ , with  $K$  being yet undefined constant matrix. The inverse of Eq. (3.1) is

$$\mathcal{P}_*(\mathcal{J}) = \int_{c-i\infty}^{c+i\infty} \frac{d\theta}{2\pi i} \exp(\theta\mathcal{J}) \mathcal{Z}(\theta; K), \quad (3.2)$$

where it is assumed that the integration contour, considered in the complex plain of  $\theta$ , goes on the right from all the singularities (poles and cuts) of  $\mathcal{Z}(\theta; K)$ . In the path integral representation GF gets the following form

$$\mathcal{Z}(\theta; K) \sim \int \mathcal{D}x \mathcal{D}p \exp \left( \int_0^T dt \left( -\frac{\theta}{2} x^* \tilde{Q} x + p^* (\partial_t x + \mu x) + \frac{1}{2} p^* V p \right) \right), \quad (3.3)$$

$$\tilde{Q} = Q + K^* R K, \quad \mu = BK - A, \quad (3.4)$$

where  $p$  is an auxiliary vector variable (momentum). Here and everywhere below we assume that, even if the dynamics was not stable before application of the control, control stabilizes it. Formally, this means that  $\mu$ , defined by Eq. (3.4), has no eigenvalues with negative real values. The "boundary" ( $F$ -dependent term) in Eq. (3.3) was ignored, assuming

that (like in the one-dimensional LQ case discussed above) it may only influence how the optimum is approached (backwards in time) but remains inessential for describing asymptotic behavior of the optimal control. This path integral is (most conveniently) evaluated by changing to the Fourier (frequency) domain, expressing pair correlation function as the frequency integral, and then relating it to the derivative of the log-GF over  $\theta$ ,

$$\langle x_i x_j \rangle = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \left( \omega^2 V^{-1} + \mu^* V^{-1} \mu + \theta \tilde{Q} \right)_{ij}^{-1}, \quad (3.5)$$

$$\frac{\partial \log \mathcal{Z}(\theta; K)}{\partial \theta} = -\frac{T}{2} \langle x^* \tilde{Q} x \rangle. \quad (3.6)$$

Here in Eqs. (3.5,3.6) the averaging is over the path integral measure described by Eq. (3.3). Further, evaluating the integral over  $\theta$ , fixing normalization,  $\mathcal{Z}(0; K) = 1$ , and using the standard formula of matrix calculus,  $d/d\theta \log \det(\theta \cdot 1 + D) = \text{tr}((\theta \cdot 1 + D)^{-1})$ , where 1 stands for the unit matrix, one arrives at the following expression

$$\log(\mathcal{Z}(\theta; K)) = -\frac{T}{2} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \log \frac{\det(\omega^2 V^{-1} + \mu^* V^{-1} \mu + \theta \tilde{Q})}{\det(\omega^2 V^{-1} + \mu^* V^{-1} \mu)}, \quad (3.7)$$

which is asymptotically exact at  $T \rightarrow \infty$ . Moreover, one can show that for any (spatially) finite system next order corrections to the rhs of Eq. (3.7) are  $O(1)$ . Note that this representation (3.7) of the log-GF, as an integral over frequency of a log-det, is similar to the relation discussed in Section 3 of [29] in the context of linking the RS-LQG control to the maximum entropy formulation of the  $H_\infty$  control. The log-det has also appeared in [38] where statistics of currents were analyzed in general non-equilibrium (off-detailed-balance) linear system.

To gain intuition let us first analyze Eq. (3.7) in the simple scalar case where the integral on the rhs can be evaluated analytically

$$\log(\mathcal{Z}(\theta; K)) = \frac{T}{2} \left( \mu - \sqrt{\mu^2 + \theta \tilde{Q} V} \right). \quad (3.8)$$

Substituting this expression into Eq. (3.2) and estimating the integral over  $\theta$  in a saddle-point approximation (justified when  $T$  is large) one arrives at the LD expression (1.8) where

$$\mathcal{S}_*(j) = \frac{V(Q + RK^2)}{16j} + \frac{(BK - A)^2 j}{V(Q + RK^2)} - \frac{BK - A}{2}. \quad (3.9)$$

The LD function is obviously convex and it is defined only for positive  $j$ . (The asterisk marks, as before, that the average and the probability are computed conditioned to yet unspecified  $K$ .)  $\mathcal{S}_*(j)$  achieves its minimum at,  $\langle j \rangle_* = -T^{-1} \partial \log \mathcal{Z} / \partial \theta|_{\theta=0} = V\tilde{Q}/(4\mu)$ , and shows linear asymptotic,  $\mathcal{S}_*(j) \approx j\mu^2/(V\tilde{Q})$ , at  $j \gg \langle j \rangle_*$ . Note, that the aforementioned asymptotic is associated with the cut-singularity in the complex  $\theta$  plane of the GF expression (3.8). Indeed, substituting Eq. (3.8) into Eq. (3.2) and shifting the integration contour to the left, thus forcing it to surround anti-clockwise the  $]-\infty; \theta_* = -\mu^2/(V\tilde{Q})]$  cut, and then estimating the integral by a small part of the contour surrounding vicinity of the cut tip at  $\theta_*$ , we arrive at the aforementioned  $j \gg \langle j \rangle_*$  asymptotic,  $\mathcal{S}(j) \approx -j\theta_*$ .

Returning back to analysis of the general formulas (3.7,3.2), one observes that even though to reconstructing  $\mathcal{S}_*(j)$  in its full integrity explicitly as a function of  $K$  does not look feasible, we can still, motivated by the scalar case analysis, make some useful general statements about both the average,  $\langle j \rangle_*$ , and the  $j \gg \langle j \rangle_*$  asymptotic of  $\mathcal{S}_*(j)$ . We will start from the latter problem.

For analysis of the tail the key object of interest is the det in Eq. (3.7) considered at zero frequency,  $\omega = 0$ . Specifically, one aims to find the zero of the determinant with the largest real value:

$$\theta_* = \max_{\theta} \text{Re}(\theta)_{\det(\mu^* V^{-1} \mu + \theta \tilde{Q})=0}. \quad (3.10)$$

Indeed, any zero (there might be many of these in the general matrix case) marks the tip of the respective cut singularity of  $\mathcal{Z}(\theta; K)$  in the complex  $\theta$ -plane. Then, the tail,  $j \gg \langle j \rangle_*$ , asymptotic of the LD function becomes,  $\mathcal{S}_*(j) = -j\theta_*$ . Note, that this linear in  $j$  estimation is valid only in the case of a finite system, when the set (spectrum) of zeros (defined by the condition in Eq. (3.10) is discrete. In the case of an infinite system, when the spectrum of zeros becomes quasi-continuous, one needs to account for the multiple zeros, as illustrated in the "string" example of Section V.



To evaluate  $\langle j \rangle_*$  (as a function of  $K$ ) in the general case one first analyzes it in the time representation. Substituting the  $u = -Kx$  ansatz with constant  $K$  in Eq. (1.1), expressing  $x(t)$  formally as an integral over time (for a given realization of the noise), substituting the result into Eq. (1.4), averaging over noise, and then taking the  $T \rightarrow \infty$  limit one arrives at

$$\langle j \rangle_* = \frac{1}{2} \int_0^\infty \text{tr} \left( V e^{-\mu^* t} \tilde{Q} e^{-\mu t} \right) dt = \frac{1}{2} \text{tr} (V \Pi) |_{\mu^* \Pi + \Pi \mu = \tilde{Q}}, \quad (3.11)$$

where the latter expression is implicit (as the condition is a matrix one, thus not resolvable explicitly in general) function of  $K$ . It is straightforward but tedious to check (introducing matrix Lagrangian multiplier for the condition in Eq. (3.11) and making variation over  $K$  and  $\Pi$ ) that optimization of Eq. (3.11) over  $K$  results in the algebraic Riccati equation equivalent to Eq. (2.2) with the first term ignored. Note that the fact that the optimal control derived from the optimization of the average cost function in the stochastic case coincides with the result of the deterministic optimization (ignoring stochasticity) is the fact very well known in the control theory. [46] The optimal value of the functional in the deterministic case saturates to a constant at  $T \rightarrow \infty$ , while in the stochastic case the average optimal cost grows with time linearly. Asymptotic convergence of the two seemingly different schemes to the same optimal control is thus an indication of the asymptotic self-consistency of the linear ansatz (1.7).

Differentiating Eq. (3.7) over  $\theta$  and then setting  $\theta$  to zero, one derives an alternative (to Eq. (3.11)) representation for the average rate of the cost function conditioned to  $K$

$$\langle j \rangle_* = \int_{-\infty}^{+\infty} \frac{d\omega}{4\pi} \text{tr} \left( (\omega^2 V^{-1} + \mu^* V^{-1} \mu)^{-1} \tilde{Q} \right). \quad (3.12)$$

Note that comparison of Eqs. (3.7, 3.11, 3.12) also allows to derive expression for the derivative of the log-GF as a time integral, and then have it presented in an implicit algebraic form

$$\begin{aligned} -T^{-1} \partial_\theta \log \mathcal{Z}(\theta; K) &= \int_{-\infty}^{+\infty} \frac{d\omega}{4\pi} \text{tr} \left( (\omega^2 V^{-1} + \mu^* V^{-1} \mu + \theta \tilde{Q})^{-1} \tilde{Q} \right) \\ &= \frac{1}{2} \text{tr} \left( \tilde{V} \Pi \right)_{\mu^* \Pi + \Pi \mu = \tilde{Q}} \end{aligned} \quad (3.13)$$

$$= \frac{1}{2} \int_0^\infty \text{tr} \left( \tilde{V} e^{-\mu^* t} \tilde{Q} e^{-\mu t} \right) dt, \quad (3.14)$$

where  $\tilde{V} = V(1 + \theta V(\mu^*)^{-1} \tilde{Q} \mu^{-1})^{-1}$ .

#### IV. OPTIMAL ASYMPTOTIC CONTROLS

In this Section we formulate the RS-LQ, TO-LQ and CC-LQ asymptotic schemes in the general vector/matrix form as an optimization over  $K$ . (Note that the asymptotic LQ scheme was already stated as a minimum of Eq. (3.11), or equivalently of Eq. (3.12) in the preceding Section.) Then we illustrate these formulations on the scalar example.

$K_{\text{RS}}$ , which is asymptotically optimal for the RS-LQ control considered at  $\theta > 0$ , is found by maximizing  $\mathcal{Z}(\theta; K)$ . Using Eq. (3.7) one derives

$$\min_K \int_{-\infty}^{+\infty} d\omega \log \left| \frac{\det(\omega^2 V^{-1} + (BK - A)^* V^{-1} (BK - A) + \theta(Q + K^* R K))}{\det(\omega^2 V^{-1} + (BK - A)^* V^{-1} (BK - A))} \right|_{\text{Re}(\lambda(BK - A)) > 0}, \quad (4.1)$$

where  $\text{Re}(\lambda(BK - A)) > 0$  denotes the stability condition ensuring that the real values of all the eigen-values of  $BK - A$  are positive. Note that constancy of the stationary RS-LQ optimal control was proven in [26], therefore making our approach self-consistent. An alternative, but obviously equivalent, formulation of the RS-LQ optimal control consists in minimizing  $-T^{-1} \partial_\theta \log \mathcal{Z}(\theta; K)$ . Going along this path and utilizing Eq. (3.13) one arrives at

$$\min_{K, \Pi} \frac{1}{2} \text{tr} \left( V(1 + \theta V(\mu^*)^{-1} \tilde{Q} \mu^{-1})^{-1} \Pi \right)_{\mu^* \Pi + \Pi \mu = Q + K^* R K}, \quad (4.2)$$

generalizing the LQ formulation stated in the preceding Section as the minimization of Eq. (3.11). Solving Eq. (4.2) is reduced to analysis of the respective generalization of the Riccati equations which can than be turned into a linear eigen-value problem described within the so-called Hamiltonian approach to the RS-LQ problem discussed in [28].

From Eq. (1.5), and assuming time-independence of the control, one can state the general asymptotic TO-LQ optimum utilizing Eqs. (3.2,3.7) as an optimization of a double integral over frequency and  $\theta$ . However, in practice one is interested to discuss the TO-LQ optimization only at sufficiently large values of the cost,  $jT$ . Using analysis of the preceding Section one derives the desired double asymptotic (valid at large  $T$  and large  $j$ ) and simpler to state expression describing  $K_{\text{TO}}$

$$\min_K \max_{\theta} \text{Re}(\theta) \quad \begin{aligned} & \text{Re}(\lambda(BK - A)) > 0 \\ & \det(\mu^* V^{-1} \mu + \theta \tilde{Q}) = 0 \end{aligned} \quad (4.3)$$

where max is over complex  $\theta$  and the optimal LD value of the PDF tail is exponential,

$$\log \mathcal{P}_{\text{TO}}(\mathcal{J}) \approx -\text{Re}(\theta_{\text{TO}}) \mathcal{J}, \quad (4.4)$$

with  $\theta_{\text{TO}}$  solving Eq. (3.10). Note that the  $\det = 0$  condition in Eq. (4.3) is reminiscent of the  $\mu$ -measure which is the key element of the robust control approach, see [35, 37] and references therein.

In the same double asymptotic (large  $T$  and large  $j$ ) regime the optimal CC-LQ control (1.6) is given by

$$\begin{aligned} \min_K \text{tr}(VII) \Big| \quad & \begin{aligned} & \text{Re}(\lambda(BK - A)) > 0 \\ & \mu^* \Pi + \Pi \mu = \tilde{Q}, \end{aligned} \\ \max_{\theta} (\text{Re}(\theta))_{\det(\mu^* V^{-1} \mu + \theta \tilde{Q})=0} & \geq \frac{\log(1/\varepsilon)}{jT}. \end{aligned} \quad (4.5)$$

Note that unlike Eqs. (4.1,4.3), Eq. (4.5) does not have valid solutions for any value of the  $\log(1/\varepsilon)/\mathcal{J}$  ratio. In fact, it is clear from Eq. (4.4) that to have a nonempty solution of Eq. (4.5) one needs to require that  $\text{Re}(\theta_{\text{TO}})jT \leq \log(1/\varepsilon)$ . Once the optimum solution is found, one estimates the LD asymptotic of the cost function PDF by an expression similar to the one given by Eq. (4.4), with TO subscript replaced by the CC one.

### A. Scalar case

In the remainder of this Section we illustrate all of the aforementioned formulas on the scalar example. In this simple case integral on the rhs of Eq. (4.1) is equal to

$$2\pi \left( \sqrt{(BK - A)^2 + V\theta(Q + RK^2)} - (BK - A) \right), \quad (4.6)$$

resulting in the following optimal value

$$K_{\theta} = \frac{A/B + \sqrt{A^2/B^2 + Q/R + QV\theta/B^2}}{1 + VR\theta/B^2}. \quad (4.7)$$

The large deviation tail of the PDF of  $j$  at a given  $K$  can be extracted from Eq. (3.9):

$$j \gg \langle j \rangle : \quad \mathcal{S}_*(j) \approx \frac{(BK - A)^2}{V(Q + RK^2)} j + o(j). \quad (4.8)$$

Optimizing the PDF over  $K$  we find two different cases depending on the sign of  $A$ . At  $A > 0$  coefficient in front of the linear in  $j$  term on the rhs of Eq. (4.8) grows monotonically with  $K$  from the  $(A/B, +\infty)$  interval. To find the optimal value of  $K$  in this case one has to take into the  $O(j)$  term thus deriving :

$$A > 0 : \quad K_{\text{TO}} = \sqrt{4AjRV}, \quad \log \mathcal{P}_{\text{TO}} \approx -\frac{B^2 jT}{RV}. \quad (4.9)$$

In the other case of  $A = -|A| < 0$  the linear coefficient in Eq. (4.8) reaches its maximum at  $K = BQ/(R|A|)$ , thus resulting in

$$A < 0 : \quad \log \mathcal{P}_{\text{TO}} \approx -\frac{jT}{RVQ} (RA^2 + B^2Q). \quad (4.10)$$

Finally, the CC-optimal formula (4.5) has no solution if  $B^2j/(RV) < c$  in the  $A > 0$  case and if  $j(RA^2 + B^2Q)/(RVQ) < c$  in the  $A < 0$  case. (Here we assume, as above, that  $\epsilon(0;T) = \exp(-cT)$ .) When  $\epsilon$  is chosen sufficiently small (i.e.  $c$  is sufficiently large), the feasibility domain in Eq. (4.5) is not empty and one distinguishes two regimes depending on how  $K_\epsilon$ , defined by

$$K_\epsilon = \frac{1}{1-\kappa} \left( \frac{A}{B} + \sqrt{\kappa \left( \frac{A^2}{B^2} + (1-\kappa) \frac{Q}{R} \right)} \right), \quad \kappa = \frac{cVR}{B^2j}, \quad (4.11)$$

compares with  $K_0$ , which is the bare LQ optimal value correspondent to  $K_\theta$  from Eq. (4.7) evaluated at  $\theta = 0$ . One derives

$$K_{CC} = \max(K_\epsilon, K_0), \quad (4.12)$$

where of the two regimes one is achieved within the interior of the optimization domain (tail constraint is not restrictive) while the other one corresponds to the tail imposed by the boundary of the domain. It is worth noting that (4.11) is valid for both signs of  $A$ .

## V. EXAMPLE OF A STRING

In this Section we discuss an explicitly solvable example of an infinite system where the set of zeros (of the determinant in the condition of Eq. (3.10)) forms a quasi-continuous spectrum. Consider a string, defined as an over-damped system of multiple bids on a line connected to each other by elastic springs of strength  $D$ , stretched by the linear force of the strength  $A$  and subject to Langevin driving:

$$\partial_t x_j = Ax_j + D(x_{j+1} + x_{j-1} - 2x_j) + Bu_j + \xi_j, \quad (5.1)$$

$$J = \frac{1}{2} \int_0^T dt \sum_j (Qx_j^2 + Ru_j^2), \quad (5.2)$$

where  $j = 1, \dots, N$ ,  $x_j$  marks position of the  $j$ -th bid of the string, and the zero-mean white-Gaussian noise is distributed as in Eq. (1.2) with  $V_{ij} = V\delta_{ij}$ .  $u_j$  in Eq. (5.1) stands for control. We are looking for a time-independent linear in  $x$  control, assuming that the control acts uniformly on all bids of the string, i.e.  $u_j = -Kx_j$ . Let us also assume that the string is periodic with the period  $N$ . Then, solution of Eq. (5.1) allows expansion in the series over spatial harmonics

$$x_j = \sum_{q=1}^N \exp(iq(j/N)) x_q, \quad (5.3)$$

with the wave vector,  $q$ , from the interval,  $-\pi < q < \pi$ , and resulting in the following separated equations for the individual harmonics

$$\partial_t x_q = Ax_q - 2D(1 - \cos q)x_q - BKx_q + \xi_q. \quad (5.4)$$

Repeating the steps leading to (3.8) one arrives at

$$\log \mathcal{Z} = \frac{T}{2} \sum_q \left( BK - A + 2D(1 - \cos q) - \sqrt{(BK - A + 2D(1 - \cos q))^2 + V(Q + RK^2)\theta} \right). \quad (5.5)$$

We choose to analyze only the most interesting regime,  $D \gg BK - A$ , when a nontrivial collective behavior emerges. Then, in the long wave-length,  $1 - \cos q \rightarrow q^2/2$ , and continuous,  $\sum_q \rightarrow (N/2\pi) \int dq$ , limits one derives

$$\log \mathcal{Z} = -\frac{TN}{3\pi} \frac{(BK - A)^{3/2}}{D^{1/2}} (1+s)^{1/4} i \left( (1 + \sqrt{1+s}) \mathcal{K} \left( \frac{1}{2} - \frac{1}{2\sqrt{1+s}} \right) - 2\mathcal{E} \left( \frac{1}{2} - \frac{1}{2\sqrt{1+s}} \right) \right), \quad (5.6)$$

$$s = \frac{V(Q + RK^2)\theta}{(BK - A)^2}, \quad (5.7)$$

where one utilizes the standard  $\mathcal{K}, \mathcal{E}$  notations for the elliptic functions.



Expression on the rhs of Eq. (5.6) shows a singularity at  $s = -1$ , coinciding with the singularity (in the complex  $\theta$  plane) observed in the scalar case at  $\theta_*$ . Substituting Eq. (5.6) into Eq. (3.2) and evaluating the integral over  $\theta$  in the saddle-point approximation one arrives at

$$\mathcal{S}_*(j) \approx \frac{\sqrt{2} N(BK - A)^{3/2}}{3\pi D^{1/2}} + \theta_* j. \quad (5.8)$$

Juxtaposing the string expression Eq. (5.8) to the scalar one Eq. (3.9) one notes different behaviors with respect to  $BK - A$ . Optimizing Eq. (5.8) over  $K$  at a given large value  $J = jT$ , one obtains the same tail expression, second formula in (4.9), however with another optimal control

$$K_{\text{str}}^{5/2} = \frac{2^{3/2} \pi A D^{1/2}}{RVTNB^{1/2}} J, \quad (5.9)$$

replacing the first formula in Eq. (4.9). Note that in the string case the optimal  $K$  scales as  $J^{2/5}$  which should be contrasted to the  $J^{1/2}$  scaling in the scalar case from Eq. (4.9).

## VI. CONCLUSIONS AND PATH FORWARD

This manuscript contributes the subject in control theory - designing control scheme with some guarantees not only on the average of the cost functions but also on fluctuations, specifically extreme fluctuations related to the tail of the cost function PDF. We consider linear, first order in time derivative, stochastic system of the Langevin type subject to minimization of a quadratic cost function and also with (chance) constraints imposed on the tail of the cost function PDF. In the stationary regime of large time, when control is sufficient to make the system stable, we reduce the stochastic dynamic problem of the "field theory" type to static optimization analysis with objectives and constraints stated in a matrix form. This type of reduction is unusual in the system lacking the fine-tuned Fluctuation Dissipation relation between relaxational and stochastic terms. On the other hand, the progress made is linked to linearity of the underlying stochastic systems which allowed, as in some problems of passive scalar turbulence [39–41] and driven linear-elastic systems [38, 42], to formally express solution for the system trajectory as an explicit function of the noise realization. Besides that, main technical ingredients, which allowed us to derive the results, consisted in making plausible assumption about the structure of the control (linear in the state variable and frozen in time), and then performing asymptotic evaluations of the cost functions statistics conditioned to the value of the cost matrix. Techniques of path integral, spectral analysis and large deviation estimations were used. We tested results on the simple scalar case and illustrated utility of the method on an exemplary high-dimensional system (1d chain of particles connected in a string).

We plan to continue exploring the interface between control theory and statistical physics addressing the following challenges.

- Computational feasibility of the main formulas of the paper, stating RS-, TO- and CC- controls in Eqs. (4.1,4.3,4.5) as static optimization problems, need to be analyzed for large systems and networks. After all main efforts in the applied control theory go into designing efficient algorithms for discovering optimal, or close to optimal, control, and we do plan to contribute this important task. Therefore, further analysis is required to answer the important practical question: if the static formulations of the newly introduced TO-QG and CC-QG controls allow computationally favorable exact or approximate expressions in terms of convex optimizations?
- We also plan to study weakly non-linear stochastic systems through a singular perturbation stochastic diagrammatic technique of the Martin-Siggia-Rose type [43]. Besides, some of the methods we used in the manuscript, especially related to large deviation analysis, are not restricted to linear systems. Our preliminary tests show that effects of the non-linearity on the PDF tail are seriously enhanced in comparison with how the same nonlinearity influences the average case control.
- It will be interesting to study TO- and CC- versions of the path-integral nonlinear control problems discussed in [13–16]. These problems, in their standard min-cost formulations, allow reduction (under some Fluctuation-Dissipation-Theorem like relations between the form of control, covariance matrix of the noise and the cost function) from the generally non-linear Hamilton-Jacobi-Bellman equations for the optimal cost function to a linear equation of a Schrödinger type.
- The effects of partial observability and noise in the observations can be easily incorporated in both TO- and CC- schemes discussed in the paper. In fact this type of generalization is standard and widespread in the control theory, where for example the LQG (Linear-Quadratic-Gaussian) control generalizes the LQ control.

- In terms of relevance to an application, this work was motivated by recent interest and discussions related to developing new optimization and control paradigms for power networks, so-called smart grids. In this application, strong fluctuations associated with loads and renewable generation, electro-mechanical control of generation, desire to make the energy production cheaper while also (and most importantly) maintaining probabilistic security limitations of the chance-constrained type - all of the above make the theoretical model discussed in this paper an ideal framework to consider. In particular, we plan to extend the approaches of [44, 45] and modify and apply the theory developed in this manuscript to design a multi-objective Chance Constrained Optimum Power Flow including better control of generation, loads and storage resources in power grids.
- We also anticipate that some of the models and results discussed in the paper are of interest for problems in statistical micro- and bio- fluidics, focusing on adjusting characteristics of individual molecules (polymers, membranes, etc) and also aimed at modifying properties of the medium (non-Newtonian flows) macroscopically. Time independent and linear nature of the control schemes discussed in the paper make them especially attractive for these applications. Natural constraints, e.g. associated with the force-field (optical or mechanical) as well as with some other physical limitations, could be incorporated into control as single- or multi-objective cost functions.

We are thankful to D. Bienstock, L. Gurvits, H.J. Kappen, K. Turitsyn and participants of the "Optimization and Control Theory for Smart Grids" project at LANL for motivating discussions and remarks. Research at LANL was carried out under the auspices of the National Nuclear Security Administration of the U.S. Department of Energy at Los Alamos National Laboratory under Contract No. DE C52-06NA25396.

- 
- [1] N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary time Series with Engineering Applications*. Wiley, New York, 1949.
  - [2] RE Kalman. Contributions to the theory of optimal control. *Bol. Soc. Mat. Mexicana*, 5:102–119, 1960.
  - [3] R. E. Kalman. A new approach to linear filtering and prediction problems. 82:35–45, 1960.
  - [4] R. L. Stratonovich. Conditional markov process theory. *Theory Prob. Appl. (USSR)*, 5:156–178, 1960.
  - [5] R. E. Kalman and R. S. Bucy. New results in lineal filtering and prediction theory. *ASME Trans., Part D (Journal of Basic Engineering)*, pages 95–108, 1961.
  - [6] R. E. Kalman. New methods of wiener filtering theory. In J. L. Bogdanoff and F. Korin, editors, *Proc. 1st Symposium, Engineering Applications of Random Function Theory and Probability*, pages 270–385. New York, Wiley, 1963.
  - [7] T. Kailath. A view of three decades of linear filtering theory. *IEEE Transactions on Information Theory*, IT-20:146–181, 1974.
  - [8] A. Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, 322:132–148, 1905.
  - [9] M. von Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, 326:756–780, 1906.
  - [10] S. Ulam. Marian smoluchowski and the theory of probabilities in physics. *American Journal of Physics*, 25:475–481, 1957.
  - [11] N. G. van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier, Amsterdam, 1992.
  - [12] C. W. Gardiner. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences*. Springer, 2004.
  - [13] H.J. Kappen. Path integrals and symmetry breaking for optimal control theory. *Journal of Statistical Mechanics: Theory and Experiment*, page P11011, 2005. Submitted.
  - [14] H.J. Kappen. Optimal control theory and the linear bellman equation. *Inference and Learning in Dynamic Models*, pages 363–387, 2011.
  - [15] J.L. van den Broek, W.A.J.J. Wiegerinck, and H.J. Kappen. Risk sensitive path integral control. *UAI*, 26:1–8, 2010. Published.
  - [16] K. Dvijotham and E. Todorov. A unifying framework for linearly solvable control. In *Proceedings of the Twenty-Seventh Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-11)*, pages 179–186, Corvallis, Oregon, 2011. AUAI Press.
  - [17] C. Jarzynski. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annual Review of Condensed Matter Physics*, 2(1):329–351, 2011.
  - [18] C. Bustamante, J. Liphardt, and F. Ritort. The nonequilibrium thermodynamics of small systems. *Physics Today*, 58:43–48, 2005.
  - [19] H. McAdams and A. Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94(3):814–819, 1997.
  - [20] T. Schmiedl and U. Seifert. Stochastic thermodynamics of chemical reaction networks. *JOURNAL OF CHEMICAL PHYSICS*, 126(4), JAN 28 2007.
  - [21] V. Chernyak, M. Chertkov, D.A. Goldberg, and K. Turitsyn. Non-equilibrium statistical physics of currents in queuing networks. *Journal of Statistical Physics*, 140:819–845, 2010.

- [22] V.Y. Chernyak, M. Chertkov, and N.A. Sinitsyn. The geometric universality of currents. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(09):P09006, 2011.
- [23] A. Ramezanpour R. Zecchina F. Altarelli, A. Braunstein. Stochastic optimization by message passing. *JSTAT*, page P11009, 2011.
- [24] J. Willems and S. Mitter. Controllability, observability, pole allocation, and state reconstruction. *IEEE Transactions on Automatic Control*, 16(6):582–595, December 1971.
- [25] M. Athans. The role and use of the stochastic linear-quadratic-Gaussian problem in control system design. *Automatic Control, IEEE Transactions on*, 16(6):529–552, 1971.
- [26] D. Jacobson. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *Automatic Control, IEEE Transactions on*, 18(2):124–131, 1973.
- [27] P. Whittle. Risk-Sensitive Linear/Quadratic/Gaussian control. *Adv. Appl. Prob.*, 13:764–777, 1981.
- [28] P. Whittle and J. Kuhn. A Hamiltonian formulation of risk-sensitive Linear/Quadratic/Gaussian control. *International Journal of Control*, 43(1):1–12, 1986.
- [29] J.C. Glover, K. and Doyle. State-space formulae for all stabilizing controllers that satisfy an H-infinity norm bound and relations to risk sensitivity. *Systems & Control Letters*, 11:167–172, 1988.
- [30] I.R. Petersen, M.R. James, and P. Dupuis. Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *Automatic Control, IEEE Transactions on*, 45(3):398–412, mar 2000.
- [31] A. Charnes, W. W. Cooper, and G. H. Symonds. Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. *Management Science*, 4:235263, 1958.
- [32] L. B. Miller and H. Wagner. Chance-constrained programming with joint constraints. *Operation Research*, 13:930945, 1965.
- [33] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.
- [34] G.F. Franklin, J.D. Powell, and A. Emami-Naeini. *Feedback Control of Dynamic Systems*. Prentice Hall, Upper Saddle River, NJ, 4th edition, 2002.
- [35] K. Zhou, J.C. Doyle, and K. Glover. *Robust Control and Optimal Control*. Prentice Hall, NJ, 1996.
- [36] J.P. Hespanha. *Linear Systems Theory*. Princeton University Press, 2009.
- [37] K. Zhou. *Essentials of Robust Control*. Prentice Hall, NJ, 1998.
- [38] K. Turitsyn, M. Chertkov, V. Y. Chernyak, and A. Puliafito. Statistics of entropy production in linearized stochastic systems. *Phys. Rev. Lett.*, 98:180603, May 2007.
- [39] M. Chertkov, G. Falkovich, I. Kolokolov, and V. Lebedev. Statistics of a passive scalar advected by a large-scale two-dimensional velocity field: Analytic solution. *Phys. Rev. E*, 51:5609–5627, Jun 1995.
- [40] M. Chertkov, G. Falkovich, and I. Kolokolov. Intermittent dissipation of a passive scalar in turbulence. *Phys. Rev. Lett.*, 80:2121–2124, Mar 1998.
- [41] G. Falkovich, K. Gawędzki, and M. Vergassola. Particles and fields in fluid turbulence. *Rev. Mod. Phys.*, 73:913–975, Nov 2001.
- [42] C. Kwon, P. Ao, and D. J. Thouless. Structure of stochastic dynamics near fixed points. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37):13029–13033, 2005.
- [43] P. C. Martin, E. D. Siggia, and H. A. Rose. Statistical dynamics of classical systems. *Phys. Rev. A*, 8:423–437, Jul 1973.
- [44] K. Dvijotham, S. Backhaus, and M. Chertkov. Operations-Based Planning for Placement and Sizing of Energy Storage in a Grid With a High Penetration of Renewables. *arxiv/1107.1382*, 2011.
- [45] S. Harnett, D. Bienstock, and M. Chertkov. Chance-constrained optimum power flows, 2012.
- [46] This fact is akin to a similar statistical physics statement naturally emerging in analysis of the linear stochastic systems driven by white-Gaussian noise and characterized by flux/current (see e.g. [38]): the average flux corresponds to the minimum of a functional quadratic in the flux.